

Agent médiateur pour la co-construction d'interprétations géoscientifiques multi-expertes

Conception, formalisation et évaluation d'un système multi-agents pour la médiation épistémique collective

Directeur de thèse	Géry Casiez, PR, HDR — Université de Lille / INRIA, équipe LOKI
Co-encadrants	Bruno Fruchard (INRIA LOKI) Cécile Gracianne (BRGM) Simon Lopez (BRGM)
Laboratoire d'accueil	Centre Inria de l'Université de Lille / Équipe LOKI, Université de Lille (partenaire PC3)
PC principal	PC3 MATCHING — Collaboration avec des systèmes intelligents

Résumé court

La production de connaissances géoscientifiques à fort enjeu sociétal (cartographie d'aléa, modélisation du sous-sol) repose sur un travail collectif multi-expert confronté à des données hétérogènes et des cadres interprétatifs différents. Cette thèse propose de concevoir, développer et évaluer un agent médiateur capable de faciliter la co-construction collective d'une interprétation en rendant visibles les hypothèses concurrentes, en maintenant la trace du raisonnement partagé, et en signalant les contradictions entre sources de données ou interprétations — sans se substituer au jugement des experts. Le sujet s'ancre dans un cas d'usage documenté (cartographie des effets de site sismique à Mayotte, BRGM/DREALM) et mobilise des approches d'argumentation computationnelle, de RAG sur corpus domaine, et d'évaluation centrée utilisateur incluant l'évaluation des effets de l'agent sur l'autonomie du raisonnement expert et la qualité de la communication entre disciplines.

Équipe d'accueil — INRIA LOKI, Université de Lille

L'équipe **LOKI** (INRIA Lille / Université de Lille) développe des recherches en interaction humain-machine, avec un focus sur les interfaces avancées, la conception d'interactions tangibles et gestuelles, et l'évaluation de systèmes interactifs complexes. Elle dispose d'une expertise reconnue en conception centrée utilisateur, en évaluation de la charge cognitive et en design de systèmes d'aide à la décision. La thèse s'inscrit dans l'axe émergent de l'équipe sur l'interaction avec des agents intelligents dans des contextes collaboratifs

Description du sujet de thèse

1.1 Contexte et scénario

La production de connaissances géoscientifiques est fondamentalement une activité collective : l'interprétation d'un phénomène géologique ou d'un aléa naturel nécessite l'intégration de sources de données dont chaque expert.e maîtrise une portion — géophysicien.nes (données électromagnétiques aéroportées, mesures sismiques de sol), géologues, data scientists, ingénieurs risques, géostatisticiens. Ces équipes multi-disciplinaires construisent une interprétation cohérente selon un **mode de collaboration épisodique** : chaque expert.e conduit ses analyses de façon autonome, puis le collectif se retrouve pour présenter ses résultats, discuter les hypothèses concurrentes et définir les tâches suivantes — avant de se disperser à nouveau. La coordination entre ces phases est aujourd'hui non structurée, et les contraintes d'agenda rendent ces moments de convergence difficiles à organiser. Dans la pratique, peu de sessions plénières sont organisées,

si bien que le chef de projet porte finalement une responsabilité forte dans la cohérence de la démarche.

Scénario de référence — Projet Mayotte (BRGM / DREALM, 2020–en cours). Pour cartographier les effets de site sismique sur l'île de Mayotte, une équipe pluridisciplinaire a produit en phase 1 — par alignement manuel entre 2 cartes géologiques, 1 333 forages, 627 mesures H/V (Rapport spectral Horizontal sur Vertical), ~800 profils MASW (Multichannel Analysis of Surface Waves) et 42 699 mesures de résistivité AEM (Airborne Electromagnetic) — un jeu de 3 271 positions d'interfaces géologiques annotées constituant la vérité terrain. La phase 2, visant à automatiser cette méthodologie à l'échelle de l'île, s'est heurtée à une limite non technique : les règles d'association entre sources hétérogènes ne peuvent pas être définies indépendamment des hypothèses géologiques que chaque expert porte implicitement. Sans espace pour expliciter et confronter ces hypothèses, le consensus sur la fusion restait inaccessible — non parce que les données étaient insuffisantes, mais parce que le raisonnement de chaque discipline restait opaque aux autres. (Gracianne et al., 2024).

1.2 Problème et objectif

L'interprétation géoscientifique est **sous-contrainte** : un même jeu de données admet plusieurs solutions plausibles, et chaque expert mobilise le cadre interprétatif de sa discipline pour naviguer dans cet espace de possibles. Dans une étude de référence portant sur 412 géoscientifiques interprétant la même image sismique synthétique, seulement 21 % ont correctement identifié le contexte tectonique (Alcalde et al., 2017). Cette divergence tient au cadre interprétatif implicite de chaque expert — ce que la littérature nomme *conceptual uncertainty* — et non à un manque de compétence.

Par ailleurs, les représentations graphiques (cartes, coupes, sections de résistivité) jouent un rôle de **boundary objects** (Star & Griesemer, 1989) : objets plastiques autour desquels des communautés aux langages différents se coordonnent sans nécessairement partager les mêmes significations. Ce statut particulier structure la collaboration, mais aussi ses frictions. Ces représentations ne sont donc pas de simples visualisations, mais des supports actifs du raisonnement sur lesquels les experts annotent, pointent et argumentent, ce qui suggère que l'interaction avec l'agent devra également s'ancrer sur ces objets graphiques.

Objectif : concevoir, développer et évaluer un agent médiateur capable d'assister un collectif d'experts géoscientifiques dans la co-construction d'une interprétation partagée, en rendant visibles les hypothèses concurrentes, en maintenant la trace du raisonnement collectif, et en signalant les contradictions entre sources ou interprétations — sans se substituer au jugement des experts ni biaiser leur raisonnement ou leur capacité à expliciter leurs hypothèses de façon autonome (Natali et al., 2025).

1.3 État de l'art

Collaboration médiatisée par IA. La littérature sur l'augmentation du raisonnement collectif distingue clairement *remplacement* (tâches routinières) et *augmentation* (décisions stratégiques à forte incertitude), et documente les risques symétriques : *automation bias*, *silent autopilot*, dérive comportementale vers la dépendance algorithmique (Cabitza et al., 2021 ; Goddard et al., 2012).

Argumentation computationnelle. Le cadre de Dung (1995) fournit une fondation mathématique pour modéliser les relations d'attaque et de défense entre hypothèses concurrentes dans un collectif. Les implémentations disponibles (*py-arg*, *Tweety*) permettent un prototypage rapide adapté à une thèse.

RAG et LLM pour les géosciences. Les approches de *Retrieval-Augmented Generation* (RAG) permettent d'ancrer les propositions d'un agent dans des connaissances domaine vérifiables, limitant les hallucinations. Des travaux récents démontrent leur faisabilité sur des données géoscientifiques structurées (GeoGPT, 2024).

Lacune identifiée. Aucun travail n'adresse spécifiquement la conception d'un agent médiateur pour un collectif multi-expert asynchrone travaillant sur des données géoscientifiques hétérogènes, ni l'impact que celui-ci aurait sur le transfert d'information entre expert·es et la communication en général lors de groupes d'études pour partager plus efficacement les connaissances de chacun·e.

1.4 Questions de recherche

- **Q1 — Interaction** : Comment concevoir les patterns d'interaction (proposition, questionnement, synthèse, signalement de contradictions) d'un agent médiateur opérant sur des objets complexes — hypothèses, cartes, modèles d'interfaces géologiques — dans un collectif asynchrone ?
- **Q2 — Formalisation** : Comment représenter et maintenir formellement l'état du raisonnement collectif (hypothèses en tension, arguments, consensus partiels) d'une session à l'autre, de manière exploitable par l'agent et lisible par les expert·es ?
- **Q3 — Évaluation** : Comment mesurer l'apport de l'agent sur la qualité de l'interprétation produite, la dynamique de collaboration, et les effets de la médiation sur l'autonomie du raisonnement expert — notamment si les expert·es conservent la capacité d'explicitier et de défendre leurs hypothèses indépendamment des propositions de l'agent ? ?

1.5 Fondements théoriques

Le sujet mobilise quatre cadres complémentaires : (1) **CSCW et théorie des boundary objects** (Star & Griesemer, 1989 ; Schmidt & Bannon, 1992) pour modéliser les objets de médiation et les frictions collaboratives ; (2) **Shared Mental Models** (Cannon-Bowers et al., 1993) pour caractériser ce que le collectif doit co-construire ; (3) **argumentation computationnelle** (Dung, 1995 ; Rahwan & Simari, 2009) pour formaliser les relations entre hypothèses ; (4) **littérature sur le deskilling et l'automation bias** (Parasuraman & Riley, 1997 ; Cabitza et al., 2021 ; Natali et al., 2025) pour cadrer l'évaluation des risques.

1.6 Approche et méthodes

La thèse s'organise en trois phases itératives sur 36 mois :

- **Phase 1 (M1–M12) — Analyse et modélisation.** Analyse rétrospective des frictions documentées dans le projet Mayotte par entretiens semi-directifs avec les experts impliqués. Définition des patterns d'interaction et du modèle formel du raisonnement collectif. Revue comparative d'approches techniques candidates (argumentation, RAG, LLM).
- **Phase 2 (M6–M24) — Prototypage.** Développement d'un prototype instrumentant un agent médiateur construit sur des composants logiciels existants, servant de support aux études d'interaction. L'environnement d'intégration sera défini au cours de la phase 1 en fonction des contraintes techniques et organisationnelles identifiées. La plateforme Galaxy, actuellement en cours de déploiement au BRGM dans le cadre de plusieurs PEPRs (IRIMA, Sous-sol, OneWater), constitue une piste candidate — son évaluation fera l'objet d'une analyse de faisabilité en début de thèse. D'autres environnements (interface dédiée, extension d'outils existants) seront considérés comme alternatives. Évaluation formative avec experts BRGM sur scénarios contrôlés issus de Mayotte. Itérations conception–évaluation.
- **Phase 3 (M18–M36) — Évaluation et généralisation.** Évaluation sommative en conditions proches du réel avec un collectif multi-expert BRGM. Test de transposabilité des principes de conception sur un deuxième cas (modélisation géologique 3D, en discussion avec S. Lopez, BRGM). Rédaction et valorisation.

1.7 Évaluation des contributions

L'évaluation combinera des approches qualitatives et quantitatives :

- **Qualité de l'interprétation produite** : comparaison entre interprétations produites avec et sans agent médiateur, évaluées par des experts indépendants et par rapport à la vérité terrain Mayotte.
- **Dynamique de collaboration** : analyse des traces d'interaction (convergence des hypothèses, fréquence des contradictions signalées, temps jusqu'au consensus).
- **Effets de la médiation sur le raisonnement expert** : évaluation de l'autonomie et de la qualité argumentative des hypothèses produites par les experts en présence de l'agent — notamment si leurs jugements restent différenciés et explicites, ou au contraire convergent

prématurément vers les suggestions de l'agent. Protocoles incluant la détection de biais d'adoption des suggestions de l'agent et combinant analyse des traces d'interaction, questionnaires d'efficacité perçue et entretiens rétrospectifs (NASA-TLX pour la charge cognitive).

- **Transposabilité** : évaluation qualitative sur un cas de généralisation (modèle géologique 3D).

Nature de la collaboration numérique

La thèse cible des activités de collaboration entre un groupe d'experts géoscientifiques (5 à 8 personnes) médiatisée par un agent intelligent. Les caractéristiques de cette collaboration sont les suivantes :

Dimension	Description
Fonction	Construction de connaissances + coordination + partage d'interprétations divergentes. L'agent assure une fonction de médiation épistémique : il explicite les désaccords implicites entre experts.
Type	Épisodique : alternance de phases de travail individuel (chaque expert analyse ses données de façon autonome) et de phases collectives (présentation des résultats, discussion des hypothèses concurrentes, définition des tâches suivantes). Ni purement synchrone, ni purement asynchrone — le défi est précisément d'assurer la continuité du raisonnement entre ces phases.
Échelle temporelle	L'enjeu central est de réduire les latences dues à la difficulté de réunir les experts en plénière : l'agent permet de structurer le raisonnement collectif entre les sessions, de sorte que chaque expert retrouve l'état des hypothèses en cours et peut y contribuer sans attendre la prochaine réunion.
Taille du groupe	5 à 8 experts issus de disciplines différentes (géophysique, géologie, data science, risques, géostatistique). Au moins 2 disciplines actives dans chaque scénario expérimental.
Espace	Sessions collectives en visioconférence ou présentiel, géographiquement distribuées. L'agent opère comme support à ces sessions et assure la continuité du raisonnement entre elles.
Rôle de l'IA	Agent visible et explicite (conforme aux exigences MATCHING). Il propose, questionne, synthétise et signale — sans décider. Son raisonnement est transparent et contestable.

Contributions à la collaboration numérique : résultats attendus et impact

La thèse vise quatre types de contributions complémentaires :

- **Empirique.** Première étude de terrain documentée sur les frictions épistémiques dans un collectif géoscientifique multi-expert asynchrone. Les données d'interaction collectées pendant les expérimentations constitueront un corpus réutilisable pour la communauté eNSEMBLE (MATCHING).
- **Théorique.** Un modèle de la collaboration médiatisée par agent dans des contextes à forte incertitude conceptuelle (*conceptual uncertainty*) et à boundary objects complexes. Ce modèle étend les travaux CSCW sur les objets de médiation à des environnements agentiques.
- **Méthodologique.** Un protocole d'évaluation de l'assistance au raisonnement collectif intégrant les effets de la médiation sur l'autonomie du raisonnement expert, applicable à

d'autres domaines (médecine, droit, analyse de données). Des guidelines de conception pour les agents médiateurs en contexte multi-expert épisodique.

- **Technique.** Un prototype d'agent médiateur construit sur des composants logiciels existants, évalué en conditions réelles avec des experts géoscientifiques.. Ce démonstrateur concret illustre la faisabilité de l'approche et sa transposabilité à d'autres domaines géoscientifiques (modèle géologique 3D France).

Impact attendu. Au-delà de la communauté géoscientifique, les résultats de la thèse sont directement pertinents pour tout contexte de collaboration multi-expert médiatisée par IA où l'incertitude conceptuelle est élevée et les experts travaillent en alternant phases autonomes et sessions de mise en commun — soit un large pan des usages professionnels de l'IA collaborative.

Positionnement dans le programme eNSEMBLE

PC3 MATCHING. Le sujet répond directement aux quatre priorités thématiques de MATCHING 2026 : (1) *modèles computationnels de collaboration médiatisée par agents* avec application à une étude de terrain ; (2) approche *multidisciplinaire* (IHM × IA agentique × géosciences) ; (3) *évaluation écologique* (experts BRGM en conditions réelles) ; (4) *gestion des compétences* via l'évaluation des effets de l'agent sur l'autonomie du raisonnement expert et la qualité de la communication entre disciplines lors de l'introduction d'un agent médiateur dans un écosystème de travail existant..

- L'agent est **visible et explicite** — son raisonnement est transparent et contestable, conformément aux exigences MATCHING.
- Le scénario implique **au moins deux collaborateurs** issus de disciplines différentes.
- L'évaluation est **écologique** (experts BRGM sur données et tâches réelles).

Lien avec les travaux eNSEMBLE existants

Le sujet s'inscrit en complémentarité naturelle avec les thèses déjà financées par eNSEMBLE en PC3 MATCHING, notamment celles portant sur les modèles computationnels de collaboration médiatisée et sur l'évaluation de l'assistance au raisonnement.

Références principales

Alcalde et al. (2017) · Cabitza et al. (2021) · Cannon-Bowers et al. (1993) · Dung (1995) · Frodeman (1995, 2014) · Goddard et al. (2012) · Gracianne et al. (2024) · Parasuraman & Riley (1997) · Rahwan & Simari (2009) · Rudwick (1976) · Schmidt & Bannon (1992) · Star & Griesemer (1989) · Natali et al. (2025)